ANALYTIC APPROXIMATION OF RATIONAL MATRIX FUNCTIONS

V.V. PELLER AND V.I. VASYUNIN

ABSTRACT. For a rational matrix function Φ with poles outside the unit circle, we estimate the degree of the unique superoptimal approximation $\mathcal{A}\Phi$ by matrix functions analytic in the unit disk. We obtain sharp estimates in the case of 2×2 matrix functions. It turns out that "generically" $\deg \mathcal{A}\Phi \leq \deg \Phi - 2$. We prove that for an arbitrary 2×2 rational function Φ, $\deg \mathcal{A}\Phi \leq 2 \deg \Phi - 3$ whenever $\deg \Phi \geq 2$. On the other hand, for $k \geq 2$, we construct a 2×2 matrix function Φ, for which $\deg \Phi = k$, while $\deg \mathcal{A}\Phi = 2k - 3$. Moreover, we conduct a detailed analysis of the situation when the inequality $\deg \mathcal{A}\Phi \leq \deg \Phi - 2$ can violate and obtain best possible results.

1. Introduction

The problem of uniform approximation of a given function on the unit circle $\mathbb{T}$ by functions analytic in the unit disk is a classical problem of approximation theory that is very important in various applications (see [Pe]). It is well known (follows from a compactness argument) that for a function φ in $L^\infty(\mathbb{T})$ there is always a best approximant $f \in H^\infty$:

$$\|\varphi - f\|_\infty = \operatorname{dist}_{L^\infty}(\varphi, H^\infty).$$

Though in general a best approximant does not have to be unique, under certain natural assumptions on φ we do have uniqueness. In particular, uniqueness holds for continuous functions φ; this was established for the first time in [Kh] and then rediscovered by several other mathematicians.

A function $\varphi \in L^\infty$ on $\mathbb{T}$ is called *badly approximable* if the zero function is a best approximant of φ. It is well known (see e.g., [Pe], Ch. 7, §5) that a nonzero continuous function is badly approximable if and only if it has constant modulus and negative winding number with respect to the origin. This criterion can be extended to functions in $QC \stackrel{\text{def}}{=} H^\infty + C \bigcap \overline{H^\infty + C}$, see [Pe], Ch. 7, §5. If φ is a nonzero rational function, then φ is badly approximable if and only if it has

constant modulus and

$$\deg \mathbb{P}_+\varphi < \deg \mathbb{P}_-\varphi \tag{1.1}$$

(this follows from the above criterion and the argument principle).

It is even more important in applications to consider the problem of analytic approximation of matrix functions on the unit circle; in systems theory scalar functions correspond to single input – single output systems, while matrix-valued functions correspond to multiple input – multiple output systems (see e.g., [Fu], [Fr], [Pe]).

It turns out, however, that the situation with best analytic approximation of matrix functions is quite different from the scalar case. Suppose that $\Phi \in L^\infty(\mathbb{M}_{m,n})$, i.e., Φ is a bounded measurable function with values in the space $\mathbb{M}_{m,n}$ of $m \times n$ matrices. A matrix function F in $H^\infty(\mathbb{M}_{m,n})$ is said to be a *best approximant of* Φ if

$$\|\Phi - F\|_{L^\infty} = \inf\{\|\Phi - Q\|_{L^\infty} : \ Q \in H^\infty(\mathbb{M}_{m,n})\}.$$

Here for a function Ψ in $L^\infty(\mathbb{M}_{m,n})$ we use the notation

$$\|\Psi\|_{L^\infty} \overset{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} \|\Psi(\zeta)\|_{\mathbb{M}_{m,n}}.$$

However, in the matrix case even for polynomial matrix functions a best approximant is almost never unique. Consider the matrix function $\Phi = \begin{pmatrix} \bar{z} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$. It is easy to see that $\|\Phi\|_{L^\infty} = 1$. On the other hand, if f is a scalar function in the unit ball of H^∞ then the function $\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & f \end{pmatrix} \in H^\infty(\mathbb{M}_{2,2})$ is a best approximant of Φ.

In the matrix case it is more natural to impose additional requirements on a best approximant to be able to choose among best approximants the "very best" one. Recall that for a matrix (or an operator on Hilbert space) A the *singular value* $s_j(A)$, $j \ge 0$, is, by definition, the distance from A to the set of matrices (operators) of rank at most j. Clearly, $s_0(A) = \|A\|$.

Definition. Given a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$, we define inductively the sets $\mathbf{\Omega}_j$, $0 \le j \le \min\{m, n\} - 1$, by

$$\mathbf{\Omega}_0 = \{F \in H^\infty(\mathbb{M}_{m,n}) : \ F \text{ minimizes} \ \ t_0 \overset{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} \|\Phi(\zeta) - F(\zeta)\|\};$$

$$\mathbf{\Omega}_j = \{F \in \Omega_{j-1} : \ F \text{ minimizes} \ \ t_j \overset{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} s_j(\Phi(\zeta) - F(\zeta))\}, \quad j > 0.$$

2

Functions in $\bigcap\limits_{k \geq 0} \mathbf{\Omega}_k = \mathbf{\Omega}_{\min\{m,n\}-1}$ are called *superoptimal approximants* of Φ by bounded analytic matrix functions. The numbers $t_j = t_j(\Phi)$ are called the *superoptimal singular values* of Φ. Note that the set $\mathbf{\Omega}_0$ consists of the best approximants of Φ.

We say that Φ is *badly approximable* if

$$\|\Phi\|_{L^\infty} = \operatorname{dist}_{L^\infty}\big(\Phi, H^\infty(\mathbb{M}_{m,n})\big)$$

and we say that Φ is *very badly approximable* if the zero function is a superoptimal approximant of Φ.

It was proved in [PY1] (see also [Pe], Ch. 14, §3) that for a continuous $m \times n$ matrix function Φ, there exists a unique superoptimal approximant F and

$$s_j\big(\Phi(\zeta) - F(\zeta)\big) = t_j(\Phi), \quad 0 \leq j \leq \min\{m,n\} - 1, \quad \text{for almost all} \quad \zeta \in \mathbb{T}.$$

For a continuous matrix function Φ, *we denote by $\mathcal{A}\Phi$ the unique superoptimal approximant of Φ*.

Note that rational matrix functions play a special role in applications in control theory; they correspond to transfer functions whose minimal realizations have finite-dimensional state spaces. Moreover, the dimension of the state space of minimal realization is equal to the (McMillan) degree $\deg \Phi$ of a matrix rational function Φ, see e.g., [Fu], [Pe] (see the definition of $\deg \Phi$ in §2). That is why it is important to obtain sharp estimates of the degree of a rational matrix function.

It is well known (see [Pe], Ch. 7, §1) that for a scalar rational function φ the following inequality holds:

$$\deg \mathcal{A}\varphi \leq \deg \varphi \quad \text{and} \quad \deg \mathcal{A}\varphi \leq \deg \varphi - 1 \quad \text{unless} \quad \varphi \in H^\infty$$

(if $\varphi \in H^\infty$, then clearly, $\mathcal{A}\varphi = \varphi$).

It was shown in [PY1] (see also [Pe], Ch. 14, §12) that if Φ is a rational matrix function with poles outside $\mathbb{T}$, then $\mathcal{A}\Phi$ is also rational. It is a very important problem to obtain sharp estimate of the degree of $\mathcal{A}\Phi$ in terms of the degree of Φ. In particular, it has been an open problem for over 10 years whether

$$\deg \mathcal{A}\Phi \leq \deg \Phi. \tag{1.2}$$

In this paper we obtain definitive results for 2×2 matrix functions. We show in §3 that (1.2) holds "generically". Moreover, generically in the nondegenerate case $\operatorname{rank}(\Phi - \mathcal{A}\Phi) = 2$ the following stronger inequality holds:

$$\deg \mathcal{A}\Phi \leq \deg \Phi - 2.$$

However, in [PY2] an example of a 2×2 rational matrix function Φ is given, for which $\operatorname{rank}(\Phi - \mathcal{A}\Phi) = 2$, but $\deg \mathcal{A}\Phi \leq \deg \Phi - 1$.

In this paper we show that actually, the degree of $\mathcal{A}\Phi$ can be greater than the degree of Φ. In §4 for $k \geq 2$, we give an example of a 2×2 matrix function Φ such that $\deg \Phi = k$, while $\deg \mathcal{A}\Phi = 2k - 3$.

3

This is the worst possible case. We show in §3 that if $k \geq 2$ and $\deg \Phi = k$, then $\deg \mathcal{A}\Phi \leq 2k - 3$. Note that the last inequality was obtained by state space methods in [LHG], but under an additional (rather complicated) assumption on Φ. Note also that the authors of [LHG] also obtained (under an additional assumption) estimates in the case of rational matrix functions of an arbitrary size.

Let us describe the results of §3 in more detail. Let

$$(\mathbb{P}_+\Phi)(z) \overset{\text{def}}{=} \sum_{j \geq 0} \hat{\Phi}(j)z^j \quad \text{and} \quad \mathbb{P}_-\Phi \overset{\text{def}}{=} \Phi - \mathbb{P}_+\Phi,$$

where $\hat{\Phi}(j)$ is the jth Fourier coefficient of the restriction of Φ to $\mathbb{T}$. Consider the matrix function Ψ defined by $\Psi = \Phi - \mathcal{A}\Phi$. Then Ψ is a very badly approximable matrix function. Clearly,

$$\mathcal{A}\mathbb{P}_-\Psi = \mathcal{A}\mathbb{P}_-\Phi = \mathcal{A}\Phi - \mathbb{P}_+\Phi = -\mathbb{P}_+\Psi$$

and it is easy to see that

$$\deg \mathcal{A}\Phi - \deg \Phi \leq \deg \mathbb{P}_+\Psi - \deg \mathbb{P}_-\Psi$$

(the degree of a rational matrix function is defined in §2). Thus the problem of estimating the degree of $\mathcal{A}\Phi$ in terms of the degree of Φ reduces to the problem of estimating the degree of $\mathbb{P}_+\Psi$ in terms of the degree of $\mathbb{P}_-\Psi$ for very badly approximable rational matrix functions Ψ.

It follows from the results of [PY1] (see also [Pe], Ch. 14, §5 and §12) that Ψ is a 2×2 very badly approximable rational matrix function if and only if the function Ψ admits on the unit circle $\mathbb{T}$ a special (thematic) factorization of the form

$$\Psi = \begin{pmatrix} \bar{w}_1 & -w_2 \\ \bar{w}_2 & w_1 \end{pmatrix} \begin{pmatrix} t_0 u_0 & \mathbf{0} \\ \mathbf{0} & t_1 u_1 \end{pmatrix} \begin{pmatrix} \bar{v}_1 & \bar{v}_2 \\ -v_2 & v_1 \end{pmatrix}, \tag{1.3}$$

where v_1, v_2, w_1, and w_2 are scalar rational functions such that $|v_1|^2 + |v_2|^2 = |w_1|^2 + |w_2|^2$ on $\mathbb{T}$, v_1 and v_2 have no common zeros in $\mathbb{D}$, and w_1 and w_2 have no common zeros in $\mathbb{D}$; $t_0 \geq t_1 \geq 0$; and u_0 and u_1 are scalar badly approximable rational functions such that $|u_0| = |u_1| = 1$ on $\mathbb{T}$.

We prove in §3 that if Ψ is nondegenerate (i.e., $t_1 > 0$) and the inequality

$$\deg \mathbb{P}_+\Psi \leq \deg \mathbb{P}_-\Psi - 2 \tag{1.4}$$

is violated, then there is a point λ in the unit disk $\mathbb{D}$ such that u_1 has a pole at λ, u_0 has a zero at λ, and both $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ have poles at $1/\bar{\lambda}$. We also show in §3 that for fixed $t_0, u_0, u_1, v_1, v_2, w_1$, and w_2, inequality (1.4) can be violated for no more than $\deg \mathbb{P}_-u_1$ numbers $t_1 \in (0, t_0]$. Moreover, we show in §3 that for Ψ given by (1.3), the sum of $[\deg \mathbb{P}_+\Psi + 2 - \deg \mathbb{P}_-\Psi]_+$ over all numbers $t_1 \in (0, t_0]$ cannot be greater than $\deg \mathbb{P}_-u_1$. Here for $x \in \mathbb{R}$, we use the notation

$$[x]_+ = \max\{x, 0\}.$$

4

In §4 we construct an example of a matrix function of the form (1.3) such that inequality (1.4) is violated for precisely $\deg \mathbb{P}_- u_1$ values of $t_1 \in (0, t_0]$. This example also shows that that the above estimate obtained in §3 is best possible.

In §2 we collect information on Hankel and Toeplitz operators, degrees of rational matrix functions, and very badly approximable matrix functions.

2. Preliminaries

2.1. Hankel and Toeplitz operators on spaces of vector functions. For a matrix function Φ in $L^\infty(\mathbb{M}_{m,n})$, the *Hankel operator*

$$H_\Phi : H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m) \stackrel{\text{def}}{=} L^2(\mathbb{C}^m) \ominus H^2(\mathbb{C}^m),$$

is defined by

$$H_\Phi f = \mathbb{P}_-(\Phi f),$$

where $\mathbb{P}_-$ is the orthogonal projection from $L^2(\mathbb{C}^m)$ onto $H^2_-(\mathbb{C}^m)$. Here $H^2(\mathbb{C}^n)$ is the Hardy space of $\mathbb{C}^n$-valued functions that is identified naturally with a subspace of $L^2(\mathbb{C}^n)$.

It is well known that

$$\|H_\Phi\| = \text{dist}_{L^\infty}\big(\Phi, H^\infty(\mathbb{M}_{m,n})\big).$$

The operator H_Φ is compact if and only if $\Phi \in \big(H^\infty + C\big)(\mathbb{M}_{m,n})$.

For $\Phi \in L^\infty(\mathbb{M}_{m,n})$, the *Toeplitz operator*

$$T_\Phi : H^2(\mathbb{C}^n) \to H^2(\mathbb{C}^m),$$

is defined by

$$T_\Phi f = \mathbb{P}_+(\Phi f),$$

where $\mathbb{P}_+$ is the orthogonal projection from $L^2(\mathbb{C}^m)$ onto $H^2(\mathbb{C}^m)$.

If Φ is a continuous square matrix function, then T_Φ is *Fredholm* (an operator T on Hilbert space is called Fredholm if its range is closed and both $\text{Ker}\, T$ and $\text{Ker}\, T^*$ are finite-dimensional) if and only if $\det \Phi$ does not vanish on $\mathbb{T}$.

We refer the reader to [Pe] for more detailed information on Hankel and Toeplitz operators on spaces of vector functions.

2.2. The degree a of rational matrix function. To define the (McMillan) degree of a rational matrix function, we define first Blaschke–Potapov products. An $n \times n$ matrix function B is called a *finite Blaschke–Potapov product* if it admits a factorization

$$B = UB_1B_2\cdots B_k, \tag{2.1}$$

5

where
$$B_j(z) = \frac{\lambda_j - z}{1 - \bar{\lambda}_j z} P_j + (I - P_j)$$

for some λ_j in $\mathbb{D}$, orthogonal projections P_j on $\mathbb{C}^n$, and a unitary matrix U. The degree of the Blaschke–Potapov product (2.1) is defined by

$$\deg B = \sum_{j=1}^{k} \operatorname{rank} P_j.$$

Let now Φ be an $m \times n$ rational function without poles in $\hat{\mathbb{C}} \setminus \operatorname{clos} \mathbb{D}$. We say that Φ has (*McMillan*) *degree* $d > 0$ (in other words, $\deg \Phi = d$) if there exists an $n \times n$ Blaschke–Potapov product B of degree d such that $\Phi B \in H^\infty(\mathbb{M}_{m,n})$, but there is no such Blaschke–Potapov product of degree $d - 1$. If Φ is a constant function, we put $\deg \Phi = 0$. It is well known (see e.g., [Pe], Ch. 2, §5) that $\deg \Phi$ is the rank of the Hankel operator $H_\Phi : H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ for rational functions Φ with poles in $\mathbb{D}$ and

$$\deg \Phi = \deg \Phi^{\mathrm{t}} \tag{2.2}$$

for an arbitrary rational matrix function Φ.

Note that it follows from (2.2) that $\deg \Phi = d > 0$ if and only if there exists a Blaschke–Potapov product B of degree d such that $B\Phi \in H^\infty(\mathbb{M}_{m,n})$, but there is no such Blaschke–Potapov product of degree $d - 1$.

If Φ is a rational matrix function without poles in $\operatorname{clos} \mathbb{D}$, we say that $\deg \Phi = \deg \Phi_\flat$, where the rational matrix function $\Phi_\flat$ is defined by $\Phi_\flat(z) = \Phi(1/z)$.

For a matrix function Φ without poles on the unit circle $\mathbb{T}$ we put

$$\deg \Phi = \deg \mathbb{P}_- \Phi + \deg \mathbb{P}_+ \Phi.$$

It is easy to see that if for $r > 0$, the function Φ_r is defined by

$$\Phi_r(z) = \Phi(rz)$$

and both Φ and Φ_r have no poles on $\mathbb{T}$, then $\deg \Phi = \deg \Phi_r$. If Φ is an arbitrary rational matrix function, we define the *degree* of Φ as

$$\deg \Phi = \deg \Phi_r,$$

where r is a positive number such that Φ_r has no poles on $\mathbb{T}$.

For a rational matrix function Φ and $\lambda \in \hat{\mathbb{C}}$ we define the (*McMillan*) *degree of* Φ *at* λ as the degree of the principal part of Ψ at λ and we use the notation $\deg_\lambda \Phi$ for the degree of Φ at λ.

Clearly,

$$\deg \Phi = \sum_{\lambda \in \hat{\mathbb{C}}} \deg_\lambda \Phi.$$

For a subset Λ of $\hat{\mathbb{C}}$, we use the notation

$$\deg_\Lambda \Phi = \sum_{\lambda \in \Lambda} \deg_\lambda \Phi.$$

It is easy to see that for $\Lambda \subset \mathbb{D}$, $\deg_\Lambda \Phi = \varkappa > 0$ if and only if there exists a Blaschke–Potapov product B of degree $\varkappa$ such that $B\Phi$ has no pole in Λ, but there is no such Blaschke–Potapov product of degree $\varkappa - 1$.

We need the following elementary facts about the degrees of rational functions.

Lemma 2.1. *Let Φ be a square rational matrix function. Then*

$$\deg_\Lambda \det \Phi \le \deg_\Lambda \Phi$$

for every $\Lambda \subset \hat{\mathbb{C}}$.

Proof. Clearly, it suffices to consider the case when $\Lambda \subset \mathbb{D}$. Let $d = \deg_\Lambda \Phi$ and let B be a Blaschke–Potapov product of degree d such that ΦB has no poles in Λ. Then $\det \Phi B = \det \Phi \det B$ has no poles in Λ. It follows immediately from the definition of Blaschke–Potapov products that $\det B$ is a scalar Blaschke product of degree d. Thus $\deg_\Lambda \det \Phi \le d$. ■

Suppose now that $\Phi = \begin{pmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{pmatrix}$ is a 2×2 matrix function. We denote by Φ^{ad} the adjugate matrix function:

$$\Phi^{\mathrm{ad}} \overset{\text{def}}{=} \begin{pmatrix} \varphi_{22} & -\varphi_{21} \\ -\varphi_{12} & \varphi_{11} \end{pmatrix}.$$

Lemma 2.2. *Let Φ be a 2×2 rational matrix function. Then*

$$\deg \Phi^{\mathrm{ad}} = \deg \Phi.$$

Proof. For the unitary matrix $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ we have $\Phi^{\mathrm{ad}} = J\Phi J^*$. It is easy to see from the definition of the degree that the degree does not change when we multiply a matrix function by a constant unitary matrix. ■

Lemma 2.3. *Suppose that Φ_1 is an $m \times n$ rational matrix function and Φ_2 is an $n \times k$ rational matrix function. Then*

$$\deg_\Lambda(\Phi_1 \Phi_2) \le \deg_\Lambda \Phi_1 + \deg_\Lambda \Phi_2$$

for every $\Lambda \subset \hat{\mathbb{C}}$.

Proof. Clearly, it suffices to prove this inequality for $\Lambda \subset \mathbb{D}$. Let us show that if Ψ is a rational function without poles on $\mathbb{T}$ and B is a finite Blaschke–Potapov product, then

$$\deg_\Lambda(\Psi B) \ge \deg_\Lambda \Psi - \deg B. \tag{2.3}$$

Let $B_\star$ be a Blaschke–Potapov product of degree $\deg_\Lambda(\Psi B)$ such that $\Psi BB_\star$ has no pole in Λ. Then $\deg(BB_\star) \geq \deg_\Lambda \Psi$. Together with the obvious equality $\deg(BB_\star) = \deg B + \deg B_\star$ this proves (2.3).

Similarly, if B is a finite Blaschke–Potapov product, then

$$\deg_\Lambda(B\Psi) \geq \deg_\Lambda \Psi - \deg B. \tag{2.4}$$

Let now B_1 be a Blaschke–Potapov product of degree $\deg_\Lambda \Phi_1$ such that $B_1\Phi_1$ has no pole in Λ and let B_2 be a Blaschke–Potapov product of degree $\deg_\Lambda \Phi_2$ such that ΦB_2 has no pole at Λ. Then by (2.3) and (2.4),

$$0 = \deg_\Lambda(B_1\Phi_1\Phi_2 B_2) \geq \deg_\Lambda(B_1\Phi_1\Phi_2) - \deg B_2 \geq \deg_\Lambda(\Phi_1\Phi_2) - \deg B_1 - \deg B_2$$

which proves the result. ∎

Remark. In fact, Lemma 2.1 in the special case of 2×2 functions is an easy consequence of Lemma 2.2:

$$(\deg_\lambda \det \Phi)^2 = \deg_\lambda((\det \Phi)I_{2\times 2}) = \deg_\lambda(\Phi^{\mathrm{ad}}\Phi) \leq \deg_\lambda \Phi^{\mathrm{ad}} \deg_\lambda \Phi = (\deg_\lambda \Phi)^2.$$

2.3. Thematic factorizations and very badly approximable matrix functions.

For simplicity, we state here results on badly approximable matrix functions in the 2×2 case. Note that all results mentioned in this subsection are also valid in the case of matrix functions of an arbitrary size; we refer the reader to [Pe] (Ch. 14) for the general case. However, in the case of an arbitrary size thematic factorizations look more complicated.

In [PY1] very badly approximable matrix functions were characterized in terms of certain special factorizations (*thematic factorizations*). In particular, it follows from the results of [PY1] that a continuous matrix function Ψ on $\mathbb{T}$ is very badly approximable if and only if it admits a factorization

$$\Psi = \begin{pmatrix} \bar{w}_1 & -w_2 \\ \bar{w}_2 & w_1 \end{pmatrix} \begin{pmatrix} t_0 u_0 & \mathbf{0} \\ \mathbf{0} & t_1 u_1 \end{pmatrix} \begin{pmatrix} \bar{v}_1 & \bar{v}_2 \\ -v_2 & v_1 \end{pmatrix}, \tag{2.5}$$

where $0 \leq t_1 \leq t_0$, u_0 and u_1 are badly approximable unimodular (i.e., $|u_0| = |u_1| = 1$ almost everywhere on $\mathbb{T}$) functions of class $QC = H^\infty + C \bigcap \overline{H^\infty + C}$; $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ are *inner and co-outer column functions* (i.e., v_1, v_2, w_1, and w_2 are H^∞ functions satisfying $|v_1|^2 + |v_2|^2 = |w_1|^2 + |w_2|^2 = 1$ almost everywhere on $\mathbb{T}$ and such that the functions v_1 and v_2 are coprime and the functions w_1 and w_2 are coprime[1]). Moreover, in this case $t_0 = t_0(\Psi)$ and $t_1 = t_1(\Psi)$ are the superoptimal singular values of Ψ. We refer the reader to [Pe], Ch. 14, §5. Note also that another (more geometric) description of very badly approximable matrix functions is given in [PT].

---

[1] Recall that H^∞ functions are called coprime if they have no common nonconstant inner divisor.

It was also shown in [PY1] that in case Ψ is nondegenerate (i.e., rank $\Psi = 2$ almost everywhere on $\mathbb{T}$), the Toeplitz operator $T_{z\Psi} : H^2(\mathbb{C}^2) \to H^2(\mathbb{C}^2)$ has dense range (see also [Pe], Ch. 14, Theorem 5.4).

It was also shown in [PY1] that if Ψ is rational, then all functions in the factorization (2.5) are rational. Moreover, the following result holds.

Lemma 2.4. *The following inequalities hold:*

$$\deg \mathbb{P}_- u_0 \leq \deg \mathbb{P}_- \Psi, \tag{2.6}$$

$$\deg \boldsymbol{v} \leq \deg \mathbb{P}_- \Psi - 1, \tag{2.7}$$

and

$$\deg \boldsymbol{w} \leq \deg \mathbb{P}_- \Psi - 1. \tag{2.8}$$

The proof of Lemma 2.4 is implicitly contained in the proof of Lemma 12.2 of Ch. 14 of[Pe]. We give here an explicit proof.

Proof of Lemma 2.4. Multiplying (2.5) on the left by $\boldsymbol{w}^{\mathrm{t}}$, we arrive at the following identity:

$$\boldsymbol{w}^{\mathrm{t}} \Psi = t_0 u_0 \boldsymbol{v}^*. \tag{2.9}$$

Applying this equality to the vector function $\boldsymbol{v}$, we obtain

$$\boldsymbol{w}^{\mathrm{t}} \Psi \boldsymbol{v} = t_0 u_0.$$

Let $\lambda \in \mathbb{D}$. Since $\boldsymbol{v}$ and $\boldsymbol{w}^{\mathrm{t}}$ are analytic in $\mathbb{D}$, we have $\deg_\lambda \boldsymbol{w}^{\mathrm{t}} = \deg_\lambda \boldsymbol{v} = 0$, and so by Lemma 2.3,

$$\deg_\lambda u_0 = \deg_\lambda (\boldsymbol{w}^{\mathrm{t}} \Psi \boldsymbol{v}) \leq \deg_\lambda \boldsymbol{w}^{\mathrm{t}} + \deg_\lambda \Psi + \deg_\lambda \boldsymbol{v} = \deg_\lambda \Psi. \tag{2.10}$$

To obtain (2.6), one has to take the sum in (2.10) over $\mathbb{D}$.

To prove (2.7), we multiply (2.9) by $1/u_0$:

$$\frac{1}{u_0} \boldsymbol{w}^{\mathrm{t}} \Psi = t_0 \boldsymbol{v}^*,$$

whence

$$\deg \boldsymbol{v} = \deg \boldsymbol{v}^* = \sum_{\lambda \in \mathbb{D}} \deg_\lambda \frac{1}{u_0} \boldsymbol{w}^{\mathrm{t}} \Psi = \sum_{\lambda \in \mathbb{D}} \deg_\lambda \frac{1}{u_0} \Psi.$$

If λ is a pole of Ψ, then by (2.10), $\deg_\lambda \Psi \geq \deg_\lambda u_0$, and since u_0 is a scalar function, we have

$$\deg_\lambda \frac{1}{u_0} \Psi \leq \deg_\lambda \Psi - \deg_\lambda u_0.$$

If λ is a regular point of Ψ, then

$$\deg_\lambda \frac{1}{u_0} \Psi = \deg_\lambda \frac{1}{u_0} = \deg_{1/\bar\lambda} u_0.$$

9

Hence, for arbitrary $\lambda \in \mathbb{D}$, we have

$$\deg_\lambda \frac{1}{u_0} \Psi \le \deg_\lambda \Psi - \deg_\lambda u_0 + \deg_{1/\overline{\lambda}} u_0.$$

Summation over $\lambda \in \mathbb{D}$ yields

$$\deg \boldsymbol{v} \le \deg \mathbb{P}_- \Psi - \deg \mathbb{P}_- u_0 + \deg \mathbb{P}_+ u_0.$$

Inequality in (2.7) follows now from (1.1). To prove (2.8), one has to apply (2.7) to the matrix function Ψ^t. $\blacksquare$

To conclude this subsection, we state one more inequality:

$$\deg \mathbb{P}_- u_1 \le \deg \mathbb{P}_- \Psi - \mu,$$

where μ is the multiplicity of the singular value 1 of the Hankel operator H_{u_0}. This is a consequence of Theorem 8.1 of Ch. 14 of [Pe]. Thus

$$\deg \mathbb{P}_- u_1 \le \deg \mathbb{P}_- \Psi - 1. \tag{2.11}$$

2.4. Singular values of H_U and H_{U^*} for unitary-valued matrix functions U. Let U be a unitary-valued $n \times n$ matrix function such that the Toeplitz operator $T_U : H^2(\mathbb{C}^n) \to H^2(\mathbb{C}^n)$ has dense range. Then the operator $H_{U^*}^* H_{U^*}$ is unitarily equivalent to the restriction of $H_U^* H_U$ to $\operatorname{Ker} T_U$ (see [Pe], Ch. 4, §4). For scalar unimodular functions this result was proved in [PK]. It is easy to see that

$$\operatorname{Ker} T_U = \{f \in H^2(\mathbb{C}^n) : \ \|H_U f\| = \|f\|\}.$$

In particular, if H_U is compact, then $\dim \operatorname{Ker} T_U < \infty$, and so

$$s_j(H_{U^*}) = s_{j+\mu}(H_U), \quad \text{where} \quad \mu = \dim \operatorname{Ker} T_U. \tag{2.12}$$

In other words, μ is the multiplicity of the singular value 1 of H_U. This means that to obtain the singular values of H_{U^*}, one has to remove from the singular values of H_U all singular values equal to 1.

Suppose now that Ψ is a continuous very badly approximable function of size $n \times n$ such that $t_j(\Psi) = 1$, $0 \le j \le n - 1$. Then Ψ is unitary-valued and the Toeplitz operator $T_{z\Psi}$ has dense range (see §2.3). Since $T_{z\Psi}$ is Fredholm (see §2.1), $\operatorname{Range} T_{z\Psi} = H^2(\mathbb{C}^n)$. In particular, for any constant vector c there exists a function $f \in H^2(\mathbb{C}^n)$ such that $T_{z\Psi} f = c$. It follows that $f \in \operatorname{Ker} T_\Psi$, and so $\dim \operatorname{Ker} T_\Psi \ge n$. Thus it follows from (2.12) that

$$\operatorname{rank} H_{\Psi^*} \le \operatorname{rank} H_\Psi - n.$$

In other words

$$\deg \mathbb{P}_+ \Psi \le \deg \mathbb{P}_- \Psi - n. \tag{2.13}$$

10

3. Upper Estimates

In this section we are going to obtain sharp estimates of the degree of the superoptimal approximant $\mathcal{A}\Phi$ of a rational 2×2 matrix function Φ in terms of the degree of Φ. As we have observed in §1, this problem reduces to the problem for a very badly approximable rational matrix function Ψ, to estimate the degree of $\mathbb{P}_+\Psi$ in terms of the degree of $\mathbb{P}_-\Psi$.

Let Ψ be a very badly approximable rational matrix function with $t_1(\Psi) \neq 0$. Without loss of generality we may assume that $t_0(\Psi) = \|H_\Psi\| = 1$. Then Ψ admits a thematic factorization

$$\Psi = \begin{pmatrix} \boldsymbol{w}^\# & \xi \end{pmatrix} \begin{pmatrix} u_0 & \boldsymbol{0} \\ \boldsymbol{0} & t_1(\Psi)u_1 \end{pmatrix} \begin{pmatrix} \boldsymbol{v}^* \\ \vartheta^{\mathrm{t}} \end{pmatrix} \tag{3.1}$$

Here u_0 and u_1 are badly approximable scalar rational functions, $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ and $\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ are rational inner and co-outer column functions, and

$$\xi = \begin{pmatrix} -w_2 \\ w_1 \end{pmatrix} \quad \text{and} \quad \vartheta = \begin{pmatrix} -v_2 \\ v_1 \end{pmatrix} \tag{3.2}$$

(see §2.3).

For a rational matrix function G we are going to use the notation

$$G^\#(\zeta) \overset{\mathrm{def}}{=} \overline{G(1/\bar{\zeta})}$$

and

$$G^* = \left(G^\#\right)^{\mathrm{t}}.$$

Clearly, $G^\#$ and G^* are rational matrix functions and

$$G^\#(\zeta) = \overline{G(\zeta)} \quad \text{for} \quad \zeta \in \mathbb{T}.$$

We consider the parametric family $\{\Psi_{[t]}\}_{t>0}$ of very badly approximable matrix functions:

$$\Psi_{[t]} = \begin{pmatrix} \boldsymbol{w}^\# & \xi \end{pmatrix} \begin{pmatrix} u_0 & \boldsymbol{0} \\ \boldsymbol{0} & tu_1 \end{pmatrix} \begin{pmatrix} \boldsymbol{v}^* \\ \vartheta^{\mathrm{t}} \end{pmatrix}. \tag{3.3}$$

As we have discussed in §2.3, $\Psi_{[t]}$ is very badly approximable for $t \leq 1$.

In this section we prove that for all but finitely many points $t \in (0, 1]$ we have

$$\deg \mathbb{P}_+\Psi_{[t]} \leq \deg \mathbb{P}_-\Psi_{[t]} - 2 \tag{3.4}$$

and for all $t \in (0, 1]$ we have

$$\deg \mathbb{P}_+\Psi_{[t]} \leq 2 \deg \mathbb{P}_-\Psi_{[t]} - 3.$$

In the case $t_0(\Psi) \neq 0$ and $t_1(\Psi) = 0$ we prove that

$$\deg \mathbb{P}_+\Psi \leq \deg \mathbb{P}_-\Psi - 1.$$

Let $\Psi = \Psi_{[t]}$ be a rational very badly approximable matrix function such that $t_1(\Psi) \neq 0$. By (3.1),

$$\Psi_{[t]} = u_0 \boldsymbol{w}^{\#}\boldsymbol{v}^* + tu_1\xi\vartheta^{\mathrm{t}}. \tag{3.5}$$

Obviously,

$$\det \begin{pmatrix} \boldsymbol{w}^{\#} & \xi \end{pmatrix}(\zeta) = \det \begin{pmatrix} \boldsymbol{v}^* \\ \vartheta^{\mathrm{t}} \end{pmatrix}(\zeta) = 1 \tag{3.6}$$

for all $\zeta \in \mathbb{T}$. Since the matrix functions in (3.6) are rational, it follows that (3.6) holds for all $\zeta \in \mathbb{C}$. Thus we have

$$\det \Psi_{[t]}(\zeta) = tu_0(\zeta)u_1(\zeta), \quad \zeta \in \mathbb{C}. \tag{3.7}$$

We start with the inequality

$$\deg \mathbb{P}_+\Psi_{[1]} \leq \deg \mathbb{P}_-\Psi_{[1]} - 2 \tag{3.8}$$

(see (2.13)).

Let $t \in (0,1]$. Suppose that both functions $u_0\boldsymbol{w}^{\#}\boldsymbol{v}^*$ and $tu_1\xi\vartheta^{\mathrm{t}}$ in (3.5) have a pole at $\lambda \in \hat{\mathbb{C}}$. We say that we have a *cancellation* at λ on the right-hand side of (3.5) if

$$\deg_\lambda(u_0\boldsymbol{w}^{\#}\boldsymbol{v}^* + tu_1\xi\vartheta^{\mathrm{t}}) < \max\{\deg_\lambda u_0\boldsymbol{w}^{\#}\boldsymbol{v}^*, \deg_\lambda u_1\xi\vartheta^{\mathrm{t}}\}.$$

Since both matrix functions $u_0\boldsymbol{w}^{\#}\boldsymbol{v}^*$ and $u_1\xi\vartheta^{\mathrm{t}}$ have rank one, for a given λ, we can have a cancellation at λ for no more than one t.

Definition. We say that t is a *nondisturbing number for the family* $\{\Phi_{[t]}\}$ (or simply a *nondisturbing number*) if for this t there is no cancellation of poles on the right-hand side of (3.5). We say that t is a *nondisturbing number for a set* $\Lambda \subset \hat{\mathbb{C}}$ if there is no cancellation of poles for $\lambda \in \Lambda$. Otherwise we say that t is a *disturbing number (disturbing number for Λ)*.

Note that the function $\boldsymbol{w}^{\#}\boldsymbol{v}^*$ is analytic in $\hat{\mathbb{C}} \setminus \operatorname{clos}\mathbb{D}$, and so if $u_0\boldsymbol{w}^{\#}\boldsymbol{v}^*$ has a pole at $\lambda \in \hat{\mathbb{C}} \setminus \operatorname{clos}\mathbb{D}$, then λ is a pole of u_0. Similarly, the function $\xi\vartheta^{\mathrm{t}}$ is analytic in $\mathbb{D}$, and so if $tu_1\xi\vartheta^{\mathrm{t}}$ has a pole at $\lambda \in \mathbb{D}$, then u_1 has a pole at λ. Thus there can be at most $\deg \mathbb{P}_+u_0$ disturbing numbers t for the exterior of the unit disk and at most $\deg \mathbb{P}_-u_1$ disturbing numbers t for the unit disk. Recall that

$$\deg \mathbb{P}_+u_0 \leq \deg \mathbb{P}_-u_0 - 1 \leq \deg \mathbb{P}_-\Psi_{[t]} - 1 \quad \text{for all} \quad t \in (0,1] \tag{3.9}$$

and

$$\deg \mathbb{P}_-u_1 \leq \deg \mathbb{P}_-\Psi_{[t]} - 1 \quad \text{for all} \quad t \in (0,1] \tag{3.10}$$

(see (1.1), (2.6), and (2.11)).

12

We are going to prove that inequality (3.4) holds for all nondisturbing numbers t for the unit disk. If 1 is a nondisturbing number for the exterior of the unit disk, then (3.4) is an immediate consequence of (3.8). In the case when 1 is a disturbing number for the exterior of the unit disk, we prove that 1 must also be a disturbing point for the unit disk and the number of cancelled poles (counted with multiplicities) outside the unit disk is at least the number of cancelled poles in the unit disk. Clearly, this would imply (3.4) for all t nondisturbing for the unit disk.

We have from (3.5)

$$\xi^* \Psi_{[t]} = t u_1 \vartheta^{\mathrm{t}} \tag{3.11}$$

and

$$\Psi_{[t]} \vartheta^{\#} = t u_1 \xi. \tag{3.12}$$

Multiplying (3.11) by the adjugate matrix $\Psi_{[t]}^{\mathrm{ad}}$ on the right and multiplying (3.12) by $\Psi_{[t]}^{\mathrm{ad}}$ on the left, we obtain

$$\left(\det \Psi_{[t]} \right) \xi^* = t u_1 \vartheta^{\mathrm{t}} \Psi_{[t]}^{\mathrm{ad}} \quad \text{and} \quad \left(\det \Psi_{[t]} \right) \vartheta^{\#} = t u_1 \Psi_{[t]}^{\mathrm{ad}} \xi.$$

It follows now from (3.7) that

$$u_0 \xi^* = \vartheta^{\mathrm{t}} \Psi_{[t]}^{\mathrm{ad}} \quad \text{and} \quad u_0 \vartheta^{\#} = \Psi_{[t]}^{\mathrm{ad}} \xi. \tag{3.13}$$

Lemma 3.1. *Let Ψ be a matrix function of the form (3.1) such that $t_1 = t_1(\Psi) \neq 0$. Suppose that u_0 has a pole of multiplicity d at λ, $|\lambda| > 1$, and $\deg_\lambda \Psi = d - l$, $l > 0$. Then u_1 has a zero at λ of multiplicity at least l, and both $\boldsymbol{v}$ and $\boldsymbol{w}$ have poles at λ of multiplicity at least l.*

Proof. By Lemma 2.1,

$$\deg_\lambda \det \Psi \leq \deg_\lambda \Psi = d - l.$$

By (3.7), u_1 has a zero at λ of multiplicity at least l. Note that $\xi^*(\lambda) \neq 0$ and $\vartheta^{\#}(\lambda) \neq 0$ (this follows from the fact that $\boldsymbol{v}$ and $\boldsymbol{w}$ are co-outer). Thus $\deg_\lambda u_0 \xi^* = \deg_\lambda u_0 \vartheta^{\#} = \deg_\lambda u_0 = d$. On the other hand, by Lemma 2.2, $\deg_\lambda \Psi^{\mathrm{ad}} = \deg_\lambda \Psi = d - l$. Hence, by (3.13), $\deg_\lambda \xi \geq l$ and $\deg_\lambda \vartheta^{\mathrm{t}} \geq l$, i.e., both $\boldsymbol{v}$ and $\boldsymbol{w}$ have poles at λ of multiplicity at least l. ■

Put

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and

$$A = \boldsymbol{w} \boldsymbol{v}^{\mathrm{t}}.$$

We have

$$\Psi_{[t]} = u_0 A^{\#} - t u_1 J A J. \tag{3.14}$$

13

Lemma 3.2. *Let $|\lambda| > 1$. Suppose that for some $t < 1$,*

$$\deg_\lambda \Psi_{[t]} - \deg_\lambda \Psi_{[1]} = l > 0. \tag{3.15}$$

Then for all $t < 1$,

$$\deg_{1/\bar\lambda} \Psi_{[t]} - \deg_{1/\bar\lambda} \Psi_{[1]} = l > 0. \tag{3.16}$$

Proof. We have already observed that since both matrix functions $u_0 \boldsymbol{w}^{\#} \boldsymbol{v}^*$ and $u_1 \xi \vartheta^{\mathrm{t}}$ have rank one, it follows from (3.5) that if (3.15) holds for some $t \in (0,1)$, then it holds for all $t \in (0,1)$. The same is true for (3.16).

Let d be the multiplicity of the pole of u_0 at λ. Since 1 is a disturbing number for the point λ, we have

$$\deg_\lambda(u_0 \boldsymbol{w}^{\#} \boldsymbol{v}^*) = \deg_\lambda(u_1 \xi \vartheta^{\mathrm{t}}).$$

Note that the vector functions $\boldsymbol{w}^{\#}$, $\boldsymbol{v}^*$ are analytic and nonzero for $|\lambda| > 1$. Therefore

$$\deg_\lambda \Psi_{[t]} = \deg_\lambda(u_0 \boldsymbol{w}^{\#} \boldsymbol{v}^*) = \deg_\lambda u_0 = d \qquad \text{for all} \qquad t \neq 1$$

and at the disturbing number the degree drops by l: $\deg_\lambda \Psi_{[1]} = d - l$.

Since for a rational function u unimodular on $\mathbb{T}$, we have $u^{\#} = 1/u$, it follows from (3.14) that

$$\Psi_{[t]}^{\#} = u_0^{\#} A - t u_1^{\#} J A^{\#} J = \frac{1}{u_0 u_1} \left(u_1 A - t u_0 J A^{\#} J \right) = -\frac{t}{u_0 u_1} J \Psi_{[1/t]} J. \tag{3.17}$$

By Lemma 3.1, the function $\frac{1}{u_0 u_1}$ at λ cannot have a zero of multiplicity greater than $d - l$. Thus it follows from (3.17) that 1 is a disturbing number for the family $\{\Psi_{[t]}^{\#}\}$ for the point λ and the multiplicity of the pole of $\Psi_{[1]}^{\#}$ at λ must also drops by l. This means that $t = 1$ is a disturbing number for the family $\{\Psi_{[t]}\}$ for the point $1/\bar\lambda$ and the multiplicity of the pole of $\Psi_{[1]}$ at $1/\bar\lambda$ must also drop by l. ■

Theorem 3.3. *For all nondisturbing numbers $t \in (0,1]$ for the unit disk inequality (3.4) holds. There can be at most $\deg \mathbb{P}_- u_1$ numbers $t \in (0,1)$, for which inequality (3.4) can be violated.*

Proof. Recall that for all nondisturbing numbers t, $\deg_\Lambda \Psi_{[t]}$ remains the same. Let us first consider the case when 1 is a nondisturbing number for the family $\{\Psi_{[t]}\}$ for the exterior of the unit disk. Then for all nondisturbing numbers t for the unit disk we have

$$\deg_{\hat{\mathbb{C}} \backslash \mathbb{D}} \Psi_{[t]} = \deg \mathbb{P}_+ \Psi_{[t]} \leq \deg \mathbb{P}_+ \Psi_{[1]},$$
$$\deg_{\mathbb{D}} \Psi_{[t]} = \deg \mathbb{P}_- \Psi_{[t]} \geq \deg \mathbb{P}_- \Psi_{[1]}.$$

Therefore inequality (3.8) yields

$$\deg \mathbb{P}_- \Psi_{[t]} - \deg \mathbb{P}_+ \Psi_{[t]} \geq \deg \mathbb{P}_- \Psi_{[1]} - \deg \mathbb{P}_+ \Psi_{[1]} \geq 2$$

for all nondisturbing numbers t for the unit disk.

14

In the case when 1 is a disturbing number for the family $\{\Psi_{[t]}\}$ for some λ, $|\lambda| > 1$, we use Lemma 3.2. Let Λ be the set of all points λ satisfying the hypotheses of Lemma 3.2, and put $\Lambda_\# = \{1/\bar\lambda \colon \lambda \in \Lambda\}$. Taking the sum in (3.15) and (3.16) over Λ yields

$$\deg_{\Lambda_\#} \Psi_{[t]} - \deg_\Lambda \Psi_{[t]} = \deg_{\Lambda_\#} \Psi_{[1]} - \deg_\Lambda \Psi_{[1]}.$$

For all $\lambda \in \hat{\mathbb{C}} \setminus (\mathbb{D} \cup \Lambda)$ the number 1 is nondisturbing, whence

$$\deg_{\hat{\mathbb{C}} \setminus (\mathbb{D} \cup \Lambda)} \Psi_{[t]} \leq \deg_{\hat{\mathbb{C}} \setminus (\mathbb{D} \cup \Lambda)} \Psi_{[1]}.$$

And finally, if t is a nondisturbing number for the unit disc, then we have

$$\begin{aligned}
\deg \mathbb{P}_- \Psi_{[t]} - \deg \mathbb{P}_+ \Psi_{[t]} &= (\deg_{\mathbb{D} \setminus \Lambda_\#} \Psi_{[t]} - \deg_{\hat{\mathbb{C}} \setminus (\mathbb{D} \cup \Lambda)} \Psi_{[t]}) + (\deg_{\Lambda_\#} \Psi_{[t]} - \deg_\Lambda \Psi_{[t]}) \\
&\geq (\deg_{\mathbb{D} \setminus \Lambda_\#} \Psi_{[1]} - \deg_{\hat{\mathbb{C}} \setminus (\mathbb{D} \cup \Lambda)} \Psi_{[1]}) + (\deg_{\Lambda_\#} \Psi_{[1]} - \deg_\Lambda \Psi_{[1]}) \\
&= \deg \mathbb{P}_- \Psi_{[1]} - \deg \mathbb{P}_+ \Psi_{[1]} \geq 2.
\end{aligned}$$

Let us estimate now the amount of disturbing numbers for the unit disc. Since the vector functions ξ and ϑ^{t} are analytic in $\mathbb{D}$, the equality

$$\Psi_{[t_1]} - \Psi_{[t_2]} = (t_1 - t_2) u_1 \xi \vartheta^{\mathrm{t}}$$

implies that the multiplicity of the pole at a point $\lambda \in \mathbb{D}$ can drop for some t only if λ is a pole of u_1. Definitely, each pole of u_1 can produce at most one disturbing number. Thus the total amount of disturbing numbers does not exceed the number of poles of u_1 in $\mathbb{D}$, what is no more than $\deg \mathbb{P}_- u_1$. ∎

The following result gives us a sharp estimate for $\deg \mathbb{P}_+ \Psi$ in the general case. The example given in the next section shows that this estimate is the best possible.

Theorem 3.4. *Let Ψ be a very badly approximable rational 2×2 matrix function such that $\deg \mathbb{P}_- \Psi = k > 0$. If $t_1(\Psi) = 0$, then*

$$\deg \mathbb{P}_+ \Psi \leq k - 1.$$

If $t_1(\Psi) \neq 0$, then

$$\deg \mathbb{P}_+ \Psi \leq 2k - 3.$$

Proof. Consider first the case when $t_1(\Psi) = 0$. Then Ψ can be represented as

$$\Psi = \begin{pmatrix} \boldsymbol{w}^{\#} & \xi \end{pmatrix} \begin{pmatrix} u_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{v}^* \\ \vartheta^{\mathrm{t}} \end{pmatrix} = u_0 \boldsymbol{w}^{\#} \boldsymbol{v}^*,$$

where u_0 is a scalar badly approximable function, $\boldsymbol{v}$ and $\boldsymbol{w}$ are inner and co-outer column functions, and ξ and ϑ are defined by (3.2). By (3.9),

$$\deg \mathbb{P}_+ u_0 < \deg \mathbb{P}_- u_0 \leq k.$$

Since $\operatorname{rank} u_0 \boldsymbol{w}^{\#} \boldsymbol{v}^* = 1$ and both $\boldsymbol{w}^{\#}$ and $\boldsymbol{v}^*$ have poles only in $\mathbb{D}$, it follows easily from Lemma 2.3 that

$$\deg \mathbb{P}_+ \Psi = \deg \mathbb{P}_+ u_0 \boldsymbol{w}^{\#} \boldsymbol{v}^* \leq \deg \mathbb{P}_+ u_0 \leq k - 1.$$

15

Suppose now that $t_1(\Psi) \neq 0$. Consider a nondisturbing number $t \in (0, 1)$. By Theorem 3.3

$$\deg \mathbb{P}_+ \Psi_{[t]} \leq \deg \mathbb{P}_- \Psi_{[t]} - 2.$$

By (3.5),

$$\Psi = \Psi_{[t_1(\Psi)]} = \Psi_{[t]} + (t_1(\Psi) - t)u_1\xi\vartheta^{\mathrm{t}}. \tag{3.18}$$

Since t is nondisturbing, we have

$$\deg \mathbb{P}_+ \Psi \leq \deg \mathbb{P}_+ \Psi_{[t]} \leq \deg \mathbb{P}_- \Psi_{[t]} - 2.$$

On the other hand, it follows from (3.18) that

$$\deg \mathbb{P}_- \Psi \geq \deg \mathbb{P}_- \Psi_{[t]} - \deg \mathbb{P}_- u_1\xi\vartheta^{\mathrm{t}} \geq \deg \mathbb{P}_- \Psi_{[t]} - \deg \mathbb{P}_- u_1$$

because both ξ and ϑ^{t} are analytic in $\mathbb{D}$. Thus

$$\deg \mathbb{P}_+ \Psi - \deg \mathbb{P}_- \Psi \leq \deg \mathbb{P}_- u_1 - 2.$$

The desired inequality follows from (3.10). ∎

Theorem 3.5. *Let Φ be a rational 2×2 matrix function such that $\deg \Phi = k > 0$ and $\mathbb{P}_- \Phi \neq \mathbf{0}$. If $t_1(\Phi) = 0$, then*

$$\deg \mathcal{A}\Phi \leq k - 1.$$

If $t_1(\Phi) \neq 0$, then

$$\deg \mathcal{A}\Phi \leq 2k - 3.$$

Proof. It suffices to apply Theorem 3.4 for the matrix function $\Psi = \Phi - \mathcal{A}\Phi$. Indeed, Theorem 3.4 yields

$$\deg(\mathbb{P}_+\Phi - \mathcal{A}\Phi) \leq \begin{cases} \deg \mathbb{P}_- \Phi - 1 & \text{if} \quad t_1(\Phi) = 0, \\ 2\deg \mathbb{P}_- \Phi - 3 & \text{if} \quad t_1(\Phi) \neq 0. \end{cases}$$

Substituting this in the inequality $\deg \mathcal{A}\Phi \leq \deg \mathbb{P}_+\Phi + \deg(\mathbb{P}_+\Phi - \mathcal{A}\Phi)$, we get

$$\deg \mathcal{A}\Phi \leq \begin{cases} \deg \mathbb{P}_+\Phi + \deg \mathbb{P}_- \Phi - 1 = \deg \Phi - 1 & \text{if} \quad t_1(\Phi) = 0, \\ \deg \mathbb{P}_+\Phi + 2\deg \mathbb{P}_- \Phi - 3 \leq 2\deg \Phi - 3 & \text{if} \quad t_1(\Phi) \neq 0. \end{cases}$$

∎

Theorem 3.4 shows that for every $t \in (0, 1)$,

$$\deg \mathbb{P}_+ \Psi_{[t]} + 2 - \deg \mathbb{P}_- \Psi_{[t]} \leq \deg \mathbb{P}_- u_1.$$

The next theorem improves this result. It shows that on the left-hand side of the above inequality we can take the sum over all $t \in (0, 1]$.

16

Theorem 3.6. *Let* $\{\Psi_{[t]}\}$, $t \in (0, 1]$, *be the family of very badly approximable matrix functions defined by* (3.3). *Then*

$$\sum_{t \in (0,1]} \left[\deg \mathbb{P}_+ \Psi_{[t]} + 2 - \deg \mathbb{P}_- \Psi_{[t]} \right]_+ \leq \deg \mathbb{P}_- u_1. \qquad (3.19)$$

Recall that $[x]_+ = \max\{x, 0\}$.

Proof. By Theorem 3.3, only finite set of disturbing points $t_1, t_2, \cdots, t_n \in (0, 1)$, $n \leq \deg \mathbb{P}_- u_1$, for the unit disk can produce a nonzero term in (3.19). Denote by Λ_j the subset of $\mathbb{D}$ such that t_j is a nondisturbing number for $\mathbb{D} \setminus \Lambda_j$ and it is a disturbing number for every point λ in Λ_j. Take any nondisturbing number t and use the same argument as in the proof of Theorem 3.4. The only difference is the following: we estimate the difference of the degrees of $\Psi_{[t_j]}$ and $\Psi_{[t]}$ not for the whole unit disk, but for each Λ_j separately. The identity

$$\Psi_{[t_j]} = \Psi_{[t]} + (t_j - t) u_1 \xi \vartheta^{\mathrm{t}}$$

implies that

$$\deg_{\Lambda_j} \Psi_{[t_j]} \geq \deg_{\Lambda_j} \Psi_{[t]} - \deg_{\Lambda_j} u_1.$$

Since t_j is nondisturbing for $\mathbb{D} \setminus \Lambda_j$, we have the following identity for this set:

$$\deg_{\mathbb{D} \setminus \Lambda_j} \Psi_{[t_j]} = \deg_{\mathbb{D} \setminus \Lambda_j} \Psi_{[t]}.$$

Thus together with the previous inequality, we obtain

$$\deg \mathbb{P}_- \Psi_{[t_j]} \geq \deg \mathbb{P}_- \Psi_{[t]} - \deg_{\Lambda_j} u_1.$$

Since t is nondisturbing, we have as before:

$$\deg \mathbb{P}_+ \Psi_{[t_j]} \leq \deg \mathbb{P}_+ \Psi_{[t]} \leq \deg \mathbb{P}_- \Psi_{[t]} - 2.$$

Therefore

$$\deg \mathbb{P}_+ \Psi_{[t_j]} + 2 - \deg \mathbb{P}_- \Psi_{[t_j]} \leq \deg_{\Lambda_j} u_1.$$

Taking the sum over j, we obtain (3.19). ∎

We will see in Theorem 4.4 that inequality (3.19) is best possible in the sense that if $\varkappa$ is an integer-valued nonnegative function on $(0, 1]$ that can be nonzero only at finitely many points, then there is a family of rational very badly approximable functions $\Psi_{[t]}$ of the form (3.3) such that

$$\left[\deg \mathbb{P}_+ \Psi_{[t]} + 2 - \deg \mathbb{P}_- \Psi_{[t]} \right]_+ = \varkappa(t)$$

and

$$\sum_{t \in (0,1]} \varkappa(t) = \deg \mathbb{P}_- u_1.$$

4. Counterexamples

In this section for $k \geq 2$, we construct an example of a very badly approximable 2×2 rational matrix function Ψ such that $\deg \mathbb{P}_- \Psi = k$, but $\deg \mathbb{P}_+ \Psi = 2k - 3$. We are going to look for Ψ in the form (3.3). It follows from Theorem 3.3 that t must be a disturbing number for the unit disk.

We also construct an example of the family $\Psi_{[t]}$ of the form (3.3) such that inequality (3.4) can be violated for several values of $t \in (0, 1]$. Actually, it can be violated for precisely $\deg \mathbb{P}_- u_1$ values of t which shows that Theorem 3.3 cannot be improved. Moreover, the same construction shows that the inequality in Theorem 3.6 cannot be improved.

The following result gives a necessary condition for the existence of a disturbing number for the unit disk that reduces the multiplicity of a pole at a given point of $\mathbb{D}$. Its proof is similar to the proof of Lemma 3.1.

Lemma 4.1. *Suppose that a point $\lambda \in \mathbb{D}$ is a pole of u_1 of multiplicity $d > 0$ and the multiplicity of the pole of Ψ at λ is $d - l$, where $l > 0$. Then u_0 has a zero at λ of multiplicity at least l and both $\boldsymbol{v}^{\#}$ and $\boldsymbol{w}^{\#}$ have a pole at λ of multiplicity at least l.*

Proof. It is easy to see that

$$\boldsymbol{w}^{\mathrm{t}} \Psi = u_0 \boldsymbol{v}^* \tag{4.1}$$

and

$$\Psi \boldsymbol{v} = u_0 \boldsymbol{w}^{\#}. \tag{4.2}$$

Since by Lemma 2.1, $\deg_\lambda \det \Psi \leq d - l$, it follows from (3.7) that u_0 has a zero at λ of multiplicity at least l.

Multiplying (4.1) on the right and (4.2) on the left by the adjugate matrix Ψ^{ad} and using (3.7), we get

$$t u_1 \boldsymbol{w}^{\mathrm{t}} = \boldsymbol{v}^* \Psi^{\mathrm{ad}} \tag{4.3}$$

and

$$t u_1 \boldsymbol{v} = \Psi^{\mathrm{ad}} \boldsymbol{w}^{\#}. \tag{4.4}$$

By Lemma 2.2, $\deg_\lambda \Psi^{\mathrm{ad}} = \deg_\lambda \Psi$. Since $\boldsymbol{v}(\lambda) \neq 0$ and $\boldsymbol{w}^{\mathrm{t}}(\lambda) \neq 0$ (their components are coprime!), the expressions in these equalities have a pole of multiplicity d at the point λ. Therefore both $\boldsymbol{v}^*$ (and hence $\boldsymbol{v}^{\#}$) and $\boldsymbol{w}^{\#}$ have poles at λ, of multiplicity at least l. ∎

We can try to construct a desired example of a function Ψ, for which all poles have multiplicity one. To construct such an example, we have to cancel $k - 1$ poles of $t u_1 \xi \vartheta^{\mathrm{t}}$ in $\mathbb{D}$. Note that $\deg \mathbb{P}_- u_1 \leq k - 1$ (see §2.3). Thus we have to cancel

all poles of $t u_1 \xi \vartheta^{\mathrm{t}}$ in $\mathbb{D}$. By Lemma 4.1, u_0 must have zeros at all poles of u_1 in $\mathbb{D}$ and both $\boldsymbol{v}^{\#}$ and $\boldsymbol{w}^{\#}$ must have poles at all poles of u_1 in $\mathbb{D}$.

Theorem 4.2. *Let $k \geq 2$ be an integer, $0 < t < 1$, and $a > 1$. Suppose that B_0, B_1, and B_2 are Blaschke products such that*

$$\deg B_0 = k - 1, \quad \deg B_1 = k, \quad \deg B_2 = k - 2,$$

and the function $1 - ta^2 B_1 B_2$ vanishes at the zeros of B_0. Let

$$\Psi = \begin{pmatrix} \frac{1}{B_0} & a \\ -a & B_0 \end{pmatrix} \begin{pmatrix} \frac{B_0}{B_1} & \mathbf{0} \\ \mathbf{0} & t\frac{B_2}{B_0} \end{pmatrix} \begin{pmatrix} \frac{1}{B_0} & a \\ -a & B_0 \end{pmatrix}. \tag{4.5}$$

Then $\deg \mathbb{P}_- \Psi = k$, while $\deg \mathbb{P}_+ \Psi = 2k - 3$.

Remark. On the right-hand side of formula (4.5) the first and the third matrix functions are not as in the definition of thematic factorizations (see (2.5)). However, if we multiply the first and the third matrix functions on the right-hand side of (4.5) by $(1 + a^2)^{1/2}$, we obtain matrix functions as in the thematic factorization (2.5). Thus the function Ψ defined by (4.5) is very badly approximable.

Proof. It is easy to verify that

$$\Psi = \begin{pmatrix} \frac{1 - ta^2 B_1 B_2}{B_0 B_1} & \frac{a}{B_1} + ta B_2 \\ -\frac{a}{B_1} - ta B_2 & B_0\left(\frac{-a^2}{B_1} + t B_2\right) \end{pmatrix}.$$

Let us first compute $\deg \mathbb{P}_- \Psi$. By the hypotheses of the theorem, $\mathbb{P}_- \Psi$ has removable singularities at the zeros of B_0. Thus $\mathbb{P}_- \Psi$ has poles only at the zeros of B_1. Hence,

$$\deg \mathbb{P}_- \Psi = \deg_{B_1^{-1}(0)} \begin{pmatrix} \frac{1}{B_0 B_1} & \frac{a}{B_1} \\ -\frac{a}{B_1} & -a^2 \frac{B_0}{B_1} \end{pmatrix}.$$

We have

$$\begin{pmatrix} \frac{1}{B_0 B_1} & \frac{a}{B_1} \\ -\frac{a}{B_1} & -a^2 \frac{B_0}{B_1} \end{pmatrix} = \frac{1}{B_1} \begin{pmatrix} \frac{1}{B_0} & a \\ -a & -a^2 B_0 \end{pmatrix}.$$

Clearly,

$$\mathrm{rank} \begin{pmatrix} \frac{1}{B_0} & a \\ -a & -a^2 B_0 \end{pmatrix} = 1,$$

and so

$$\deg \mathbb{P}_- \Psi = \deg B_1 = k.$$

Let us now compute $\deg \mathbb{P}_+ \Psi$. Clearly, $\mathbb{P}_+ \Psi$ can have poles only at the poles of B_2 and the poles of B_0.

Let us first show that there can be no cancellation of poles at the poles of B_2 or B_0. The only possibility of cancellation is to cancel poles of B_0 of the lower right

entry which can be done if $-\frac{a^2}{B_1} + tB_2$ vanishes at poles of B_0. Let ζ be a zero of B_0 and suppose that

$$\frac{a^2}{B_1(1/\bar{\zeta})} = tB_2(1/\bar{\zeta}).$$

It follows that

$$a^2 B_1(\zeta) B_2(\zeta) = t.$$

Since $t < 1$, this contradicts the assumption that $ta^2 B_1(\zeta) B_2(\zeta) = 1$.

First, we compute the contribution of the poles of B_0. Clearly, only the lower right entry of $\mathbb{P}_+\Psi$ has poles at poles of B_0. Obviously,

$$\deg_{B_0^{-1}(\infty)} \Psi = k - 1.$$

Let us compute the contribution of the poles of B_2. We have

$$\deg_{B_2^{-1}(\infty)} \Psi = \deg_{B_2^{-1}(\infty)} \begin{pmatrix} \frac{-ta^2 B_2}{B_0} & taB_2 \\ -taB_2 & tB_2B_0 \end{pmatrix} = \deg_{B_2^{-1}(\infty)} B_2 \begin{pmatrix} \frac{-ta^2}{B_0} & ta \\ -ta & tB_0 \end{pmatrix}.$$

Clearly,

$$\mathrm{rank} \begin{pmatrix} \frac{-ta^2}{B_0} & ta \\ -ta & tB_0 \end{pmatrix} = 1,$$

and so the contribution of the poles of B_2 is $\deg B_2 = k - 2$. Thus

$$\deg \mathbb{P}_+\Psi = k - 1 + k - 2 = 2k - 3. \qquad \blacksquare$$

Corollary 4.3. *Let Ψ be the function defined in Theorem 4.2 and let $\Phi = \mathbb{P}_-\Psi$. Then $\deg \Phi = k$ and $\deg \mathcal{A}\Phi = 2k - 3$.*

Finally, we can slightly modify the construction in Theorem 4.2 to get a family $\Psi_{[t]}$ of the form (3.3) such that inequality (3.4) is violated for several different values of $t \in (0, 1)$.

The following result shows that roughly speaking, inequality (3.19) is the only constraint for the numbers $\left[\deg \mathbb{P}_+\Psi_{[t]} + 2 - \deg \mathbb{P}_-\Psi_{[t]} \right]_+$.

Theorem 4.4. *Let k, a, B_0, B_1, and B_2 be as in Theorem 4.2. Suppose that $t_1, \cdots, t_d$ are distinct points in $(0, 1)$, $\Delta_1, \cdots, \Delta_d$ is a partition of the zero set of B_0, and $\varkappa_j$ is the number of zeros of B_0 in Δ_j (counted with multiplicities). Assume that*

$$1 - t_j a^2 B_1 B_2 \quad \text{vanishes at the points of} \quad \Delta_j, \quad 1 \leq j \leq d. \qquad (4.6)$$

If the very badly approximable function $\Psi_{[t]}$ is defined by

$$\Psi_{[t]} = \begin{pmatrix} \frac{1}{B_0} & a \\ -a & B_0 \end{pmatrix} \begin{pmatrix} \frac{B_0}{B_1} & \mathbf{0} \\ \mathbf{0} & t\frac{B_2}{B_0} \end{pmatrix} \begin{pmatrix} \frac{1}{B_0} & a \\ -a & B_0 \end{pmatrix}, \quad 0 < t \leq 1,$$

then inequality (3.4) is violated precisely for $t = t_1, \cdots, t_d$ *and*

$$\deg \mathbb{P}_+ \Psi_{[t_j]} = \deg \mathbb{P}_- \Psi_{[t_j]} - 2 + \varkappa_j.$$

The proof of Theorem 4.4 is exactly the same as the proof of Theorem 4.2.

In particular, if each Δ_j is a singleton, inequality (3.4) is violated for precisely $\deg \mathbb{P}_- u_1$ values of t.

Remark. Let us fix the zeros of B_0. Condition (4.6) means that the Blaschke product $B_1 B_2$ interpolates at the points of Δ_j the value $(a^2 t_j)^{-1}$. It is well known (see e.g., [Pe], Ch. 7, § 1) that if a is sufficiently large, it is possible to find a Blaschke product of a given degree greater than or equal to k that solves this interpolation problem. Thus if a is sufficiently large, it is possible to find B_1 and B_2 satisfying the hypotheses of the theorem.

References

[Fr] B.A. Francis, *A Course in H^∞ Control Theory*, Lecture Notes in Control and Information Sciences **88**, Springer-Verlag, Berlin, 1986.

[Fu] P.A. Fuhrmann, *Linear Systems and Operators in Hilbert Space*, McGraw-Hill, New York, 1981.

[Kh] S. Khavinson, *On some extremal problems of the theory of analytic functions*, Uchen. Zapiski Mosk. Universiteta, Matem. **144**:4 (1951), 133–143. English Translation: Amer. Math. Soc. Translations (2) **32** (1963), 139–154.

[LHG] D.J.N. Limebeer, G.D. Halikias, and K. Glover, *State-space algorithm for the computation of superoptimal matrix interpolating functions*, Int. J. Control **50** (1989), 2431–2466.

[Pe] V.V. Peller, *Hankel operators and their applications*, Springer-Verlag, New York, 2003.

[PK] V. V. Peller and S. V. Khruschëv, *Hankel operators, best approximation and stationary Gaussian processes*, Uspekhi Mat. Nauk, **37**:1 (1982), 53-124. English Transl.: Russian Math. Survey., **37**:1 (1982) 61-144.

[PT] V.V. Peller and S.R. Treil, *Very badly approximable matrix functions*, Selecta Mathematica, **11** (2005), 127–154.

[PY1] V.V. Peller and N.J. Young, *Superoptimal analytic approximations of matrix functions*, J. Funct. Anal. **120** (1994), 300–343.

[PY2] V.V. Peller and N.J. Young, *Construction of superoptimal approximation*, Math. Control Signals Systems **8** (1995), 118–137.

V.V. Peller
Department of Mathematics
Michigan State University
East Lansing, Michigan 48824
USA

V.I. Vasyunin
St.Petersburg Branch
Steklov Institute of Mathematics
Fontanka 27, 191011 St-Petersburg
Russia